**QUEEN'S
UNIVERSITY
BELFAST**

Reimagining the Internet with Edge Computing

Varghese, B., & Bonomi, F. (2021). Reimagining the Internet with Edge Computing. *ITNOW*, 63(1).
<https://doi.org/10.1093/itnow/bwab007>

Published in:
ITNOW

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2021 BCS, The Chartered Institute for IT

This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Reimagining the Internet with Edge Computing

Blesson Varghese, Associate Professor in computer science at Queen's University Belfast and Flavio Bonomi, Founder of Nebbiolo Technologies and Board Technology Adviser for LYNX, explore how edge computing could foster a more ethical and fairer internet.

The underlying model adopted by many Internet applications relies on services offered by centralised, power hungry and geographically distant clouds. Typically, a user interacts with a device to generate data that is sent to the cloud for processing and storage. This model works reasonably well for many applications we rely on daily. However, the rapid expansion seen in the Internet, the need for real-time data processing in futuristic applications, and the global shift in attitudes towards a more ethical and sustainable Internet have all challenged the current working model.

There are primarily four arguments that favour a change to the underlying Internet model. They are the 4 P's:

1. **Proliferation** - can the existing network infrastructure cope with the billions of users and their manifold devices that need to be connected?
2. **Proximity** - are users sufficiently close to data processing centres to achieve the real-time responses required by critical applications?
3. **Privacy** - can we control our data and the subsequent choices and actions arising from processing it?
4. **Power** - is it sustainable to rely on power hungry cloud data centres?

Leveraging compute capabilities that are placed at the edge of the network for processing data appeals to all of the above arguments. Firstly, (pre)processing data at the edge would reduce data traffic beyond the edge, thus reducing the ingress bandwidth demand to the core network.

Secondly, with compute locations closer to the user, the communication latencies can be significantly reduced when compared to the cloud, thus making applications responsive enough for real-time use.

Thirdly, localised processing of data acts as a privacy firewall to ensure selective release of user data beyond one's familiar geographic boundary and legal jurisdiction.

And finally, fewer resources will be hosted in many more edge locations, thus offering an energy and environmentally friendly alternative to traditional data centres.

Understanding the edge advantage

Edge computing is generally understood to be the integrated use of resources located in a cloud data centre and along the entire continuum towards the edge of the network for distributing Internet applications to (pre)process data and address the challenges posed above. The edge may refer either to infrastructure or user edge.

The infrastructure edge refers to edge data centres for example that are deployed on the telecom operator side of the last mile network. The user edge refers to resources, such as end-user devices or home routers and gateways that are located on the user side of the last mile network.

The origins of edge

The initial ideas relevant to edge computing were envisioned over a decade ago in a 2009 article entitled, “The Case for VM-Based Cloudlets in Mobile Computing” by M.

Satyanarayanan et al. Fog computing is an alternate term referring to the same concept that was presented by Cisco in 2012 in an article entitled, “Fog Computing and its Role in the Internet of Things” by F. Bonomi (co-author of this article) et al.

Two categories of applications have been identified to benefit from edge computing. The first is referred to as edge enhanced (or edge accelerated). These applications may be native to the cloud or the user device and will achieve a performance or functionality gain when selected services of the application are moved to the edge.

The more important class is edge native applications – they cannot emerge in the real world without the use of the edge. Illustrative examples include those that:

- Augment human cognition, for example providing real-time cognitive assistance to the elderly or those with neurodegenerative conditions using wearables.
- Perform live video analytics from a large array of cameras, including real-time denaturing for privacy.
- Use machine learning for predictive safety in driverless cars and predictive quality control in manufacturing.

Edge computing is still in its infancy, and some have looked at it with scepticism. However, there are a number of market trends that assure that edge computing is here to stay, not merely a fad that will dissipate over time.

Looking over the horizon

Substantial investments have been made by major telecom and internet providers, hardware manufacturers and vendors and technology providers in the UK and worldwide. This is evidenced by the wide variety of edge-enabled gateways and routers and modular data centres suitable for the infrastructure edge that is emerging. The strategic direction of the ETSI (European Telecommunications Standards Institute) mobile edge computing initiative is to bring compute and storage to base stations. The design and development of novel processors suited for the edge and edge specific workloads is another indicator.

However, there are many more miles for the edge computing concept to go before it can be adopted in the mainstream as an alternate model for the Internet. A number of technical challenges combined with a mixed bag of legal, business and geo-political barriers need to be overcome. We examine some of these in the remainder of this article. Although they do indeed pose challenges, they are more productively viewed as unique opportunities in the technology landscape for the creation of a more ethics-oriented Internet that will occupy engineers, programmers and scientists for the next couple of decades.

Convergence of IT and telecoms isn't easy

The technical and business case for edge computing is clear – the value proposition is to bring cloud-like applications and services (more IT and software driven) onto the network edge (more telecoms and hardware driven). The realisation of this aim requires IT and telecom providers who have different specialisations – to generalise, one with software and the other with hardware – to join forces in delivering edge computing infrastructure, platforms and services to the masses. However, both sets of providers have conflicting

business ideals; the IT enterprise may want to have monopoly over data so that it is processed in the cloud, whereas telecom providers may favour decentralisation of compute and storage away from cloud data centres to the network edge.

Performance and Safety

Ensuring performance and safety guarantees have proven to be nearly impossible when using IT software. Although computing in the IT world is fast, it is not deterministic; it is unpredictable. Hardware and software guarantees are essential for the reliability and safety of the type of critical applications envisioned by edge native applications. Hardware and software in the embedded world offer real-time and safety guarantees, but such guarantees cannot be achieved with typical IT hardware and software. Consider for example, certifying systems against safety or security failures. This is a complex task to achieve when adopting typical IT computing and software and is aggravated by the fact that software is fragmented and relies on the use of multiple languages even in the same software stack. For this reason, we are now witnessing a progressive convergence of technologies and approaches developed for embedded systems into edge computing, to make it viable for the support of mission critical applications (<https://www.lynx.com/executive-blog/from-mission-critical-embedded-systems-to-mission-critical-edge>).

Accountability as a key offering

Many users worldwide have lost their trust in the Internet due to the scores of scandals that have arisen from corporations that have misused their monopoly over user data in the cloud. However, the edge potentially offers a distinct and unique opportunity to restore this trust - decentralised data processing combined with a user's ability, by interacting with the underlying infrastructure and algorithms, to control their data and the actions and choices originating from processing it. This is challenging as multiple stakeholders, namely network, edge, application and cloud providers and end users will need to interact. Some of the answers to these challenging problems could lie in innovative and scalable data ledger solutions that will allow users to track their data in the network and interact with the elements operating on them.

Democratisation of the edge

The edge offering can truly be a success if more than a select few can monopolise it. It needs to become a citizens' endeavour. The role of a unified programming model and language to express edge applications and associated technologies for inviting broader participation among different stakeholders is immense. This may take the form of a sufficiently abstract model underpinned by automated approaches that can seamlessly switch between the contexts of different technologies for the less experienced and a more fine-grain model offering control of the underlying resources for the more savvy.

Many other technological concerns

The network fabric for seamlessly connecting the enterprise, the cloud and the edge need to mature. Stronger ways of partitioning the network and compute resources than those available today are required for efficient resource management. Since the edge will open up a hugely heterogeneous space of different generations of technologies, hardware and architectures on the infrastructure, a unified model for deploying applications on them is required. Current application orchestrators that are based on containers are limited when operating across different ownerships and hardware architectures. Software engineering

practices, for example DevOps, will need to substantially evolve, offering more automated tools to embrace the convergence of telecoms and IT.

The golden era of the Internet is yet to be realised as numerous challenges need to be tackled. Undeniably, it is an exciting prospect that edge computing holds the key to unlocking a lot of the potential of the future Internet.

Further Reading:

- B. Varghese et al. "Challenges and Opportunities in Edge Computing," IEEE International Conference on Smart Cloud, 2016, pp. 20-26.
- C. -H. Hong and B. Varghese, "Resource Management in Fog/Edge Computing: A Survey on Architectures, Infrastructure, and Algorithms," ACM Computing Surveys, Vol. 52, Issue 5, Article No. 97, 2019.